The Dynamics, Structure, and Conformational Free Energy of Proline-Containing Antifreeze Glycoprotein

Dat H. Nguyen,*†¶ Michael E. Colvin,† Yin Yeh,‡ Robert E. Feeney,§ and William H. Fink*
Departments of *Chemistry, ‡Applied Science, §Food Science and Technology, and ¶Computer Science, University of California, Davis, California 95616; and †Division of Computational and Systems Biology, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, Livermore, California 94550 USA

ABSTRACT Recent NMR studies of the solution structure of the 14-amino acid antifreeze glycoprotein AFGP-8 have concluded that the molecule lacks long-range order. The implication that an apparently unstructured molecule can still have a very precise function as a freezing inhibitor seems startling at first consideration. To gain insight into the nature of conformations and motions in AFGP-8, we have undertaken molecular dynamics simulations augmented with free energy calculations using a continuum solvation model. Starting from 10 different NMR structures, 20 ns of dynamics of AFGP were explored. The dynamics show that AFGP structure is composed of four segments, joined by very flexible pivots positioned at alanine 5, 8, and 11. The dynamics also show that the presence of prolines in this small AFGP structure facilitates the adoption of the poly-proline II structure as its overall conformation, although AFGP does adopt other conformations during the course of dynamics as well. The free energies calculated using a continuum solvation model show that the lowest free energy conformations, while being energetically equal, are drastically different in conformations. In other words, this AFGP molecule has many structurally distinct and energetically equal minima in its energy landscape. In addition, conformational, energetic, and hydrogen bond analyses suggest that the intramolecular hydrogen bonds between the N-acetyl group and the protein backbone are an important integral part of the overall stability of the AFGP molecule. The relevance of these findings to the mechanism of freezing inhibition is discussed.

INTRODUCTION

Many arctic and Antarctic fish species synthesize high concentrations of glycoproteins (~35 mg/l) in their body tissue to prevent freezing while in the subzero degree Celsius environment (Yeh and Feeney, 1996; Ben, 2001). These glycoproteins, called antifreeze glycoproteins, belong to the class of biological antifreezes, which consists of both antifreeze proteins (AFPs) and antifreeze glycoproteins (AFGPs). Members of the AFGP family are very similar in chemical composition and there are eight of them characterized to date (DeVries et al., 1970, 1971; Komatsu et al., 1970; DeVries, 1971a, b; Duman and DeVries, 1972; Feeney and Hofmann, 1973; Feeney, 1974). They all contain a repeated tripeptide sequence of alanine-alanine-threonine with a disaccharide (3-*O*-(β-D-galactosyl)-D-*N*-acetylgalactosamine) bonded to the β-oxygen of the threonine residue. They differ from one to another mainly in the number of repeated tripeptide units, although for some small AFGPs, one or two alanines are substituted by proline and by arginine.

AFPs (Yeh and Feeney, 1996; Harding et al., 1999; Fletcher et al., 2001), however, are very diverse in both chemical composition and structure. Members in this family are classified further into four types (types 1–4), each possessing completely different structures, and do not bear carbohydrate groups (Duman and DeVries, 1974, 1976; Hew et al., 1985; Ng et al., 1986; Slaughter et al., 1981; Ng and Hew, 1992; Jia et al., 1995; DeLuca et al., 1996; Sönnichsen et al., 1996; Deng et al., 1997; Liou et al., 2000). AFPs are found in the winter flounder, sculpin family, sea raven and ocean pout, plants, and insects.

Biological antifreezes are known macroscopically not only to prevent the growth of the incipient ice crystals within a certain range of temperature of supercooling, but also to interfere with the growth morphology of ice crystal and, in a noncolligative manner, to exhibit thermal hysteresis, which lowers the freezing temperature without affecting the melting temperature. Furthermore, the thermal hysteresis is additive to the colligative effect. The hypothesis (see Yeh and Feeney, 1996 for complete review of the postulates) on the mechanism of action of biological antifreezes is that the irreversible binding of biological antifreeze molecules to the ice surface through the hydrogen-bonding network is responsible for the hysteresis as governed by the Kelvin relation. However, work by Kuroda (1991) has raised doubt about the irreversible binding aspect of the hypothesis. Recent work using site-directed mutagenesis of winter flounder AFP (Chao et al., 1997; Haymet at al., 1998, 1999, 2001) has shown that hydrogen bonding between this molecule and the ice may not be responsible for the thermal hysteresis. Thus, the exact molecular mechanism of action of biological antifreezes at the molecular level still remains to be elucidated.

The AFGP-8, the focus of this work, is the smallest AFGP within the AFGP family, with only 14 amino acid residues with the following sequence:

ALA-ALA-THG-ALA-ALA-THG-PRO-ALA-THG-ALA-ALA-THG-PRO-ALA

Submitted July 27, 2001, and accepted for publication March 6, 2002.

Address reprint requests to Dr. William H. Fink, Department of Chemistry, University of California, One Shields Avenue, Davis, California 95616. Tel.: 530-752-0935; Fax: 530-752-8995; E-mail: fink@chem.ucdavis.edu.

0006-3495/02/06/2892/14 $2.00

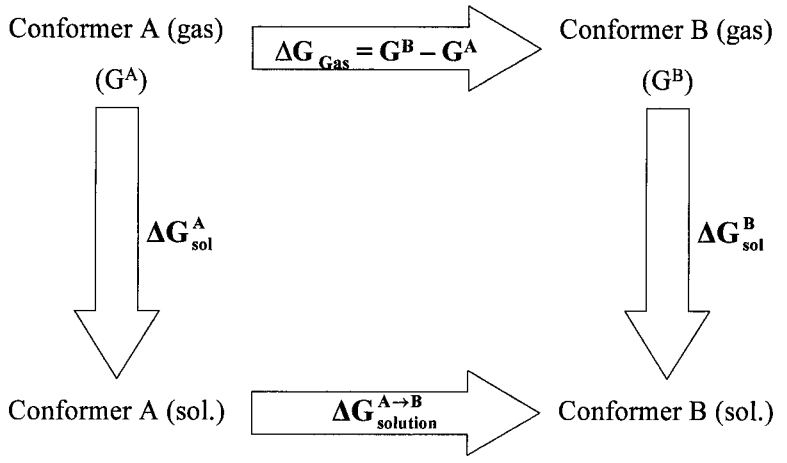

FIGURE 1 The thermodynamic cycle for comparing free energies of two conformers A and B in solution

where THG is the threonine residue linked to the disaccharide 3-*O*-(β-D-galactosyl)-D-*N*-acetylgalactosamine at the β-oxygen. This particular member of the AFGP family is mostly found in the Antarctic cod, *Pathogenia borchgrevinki* (earlier classified as *Trematomus borchgrevinki*). An aqueous solution of AFGP-8 exhibits freezing inhibition if the supercooling is modest, AFGP-8 works cooperatively with AFGP-1 to AFGP-5 under more extreme conditions, and works by itself as well, especially in recrystallization inhibitors. Furthermore, AFGP-8 is the only AFGP found in the tissue samples outside the blood stream (Burcham et al., 1986; Mulvihill et al., 1980). The x-ray structure for AFGP-8 is not yet known, but experimental work on the structural elucidation of this AFGP using a variety of experimental techniques, such as nuclear magnetic resonance (NMR) and ultra-ultraviolet circular dichroism (CD) (Bush et al., 1981, 1984; Rao and Bush, 1987; Filira et al., 1990; Dill et al., 1992; Mimura et al., 1992; Lane et al., 1998) have yielded rather confusing and conflicting messages, from the standpoint of structural biology. NMR work by Bush et al. and Lane et al. suggested that, because NMR spectra show an absence of long-range coupling, AFGP-8 does not have well-defined structure at all. Bush et al. interpreted their results in such a way that the structure of AFGP-8 is a flexible rod conformation with locally defined structure and sufficient segmental mobility to destroy any long-range order. Lane et al. observed the same results 15 years later with more advanced NMR technology; they showed that AFGP-8 does not seem to adopt a rigid, well-defined structure at all even at the near-freezing temperature, as commonly assumed for any class of proteins from their primary amino acid sequences. NMR work by Mimura et al. suggested the existence of hydrogen bonding between the N-acetyl groups of the sugars to the protein backbone. In addition, they also suggested that the role of proline residue found in some small AFGPs (including AFGP-8) is to stabilize structures of the AFGPs that resemble polyproline II structures. However, Dill et al. later interpreted the intramolecular hydrogen-bonding motif differently based on the weakness of the signal in an NMR proton exchange experiment.

For the purpose of elucidating the atomistic details of the structural dynamics and determining the native state of AFGP-8, in this study we present multi-nanosecond dynamics simulations of AFGP-8 in an explicit water environment, starting with NMR structures resulting from the work of Lane and co-workers. Two major results are reported in this paper: the structural dynamics of the AFGP-8 molecule and the determination of the lowest free-energy structures of the AFGP-8 as determined from calculation of the relaxation free energies of conformations using a continuum solvent model. From the first result, we show which conformations AFGP-8 may likely adopt during the course of the dynamics and from the second, we will present structures that have lowest conformational free energy. In addition to these two major results, we also examined the intramolecular hydrogen bond between the N-acetyl group to the protein backbone to shed light on the differing interpretations of Mimura et al. and Dill et al. Finally, we present an algorithm for determining the low-energy structures of highly flexible proteins from molecular dynamics (MD) simulations.

CONFORMATIONAL FREE ENERGY EVALUATION

The basic procedure we used for computing the free energy difference between two states has been well established (Smith and Honig, 1994; Yang et al., 1996; Bashford et al., 1997; Demchuk et al., 1997; Srinivasan et al., 1998; Jayaram et al., 1998). Here, we summarize the essence of this method.

The free energy difference between two distinct conformers A and B of a molecule in solution can be determined using the thermodynamic cycle shown in Fig. 1. In this figure, the free energy associated with the conformational difference in solution is given by:

$$\Delta G^{A\to B}_{solution} = (\langle \Delta G^{B}_{sol} \rangle + \langle G^{B} \rangle) - (\langle \Delta G^{A}_{sol} \rangle + \langle G^{A} \rangle) \quad (1)$$

where $\langle \Delta G^i_{sol} \rangle$ and $\langle G^i \rangle$ are the averages of solvation energy and the conformational energy components of the free energy for the conformer i, respectively. The solvation energy term can be estimated as follows:

$$\langle \Delta G^i_{sol} \rangle = \langle \Delta G^i_{elec} \rangle + \langle \Delta G^i_{np} \rangle \quad (2)$$

The first term on the right-hand side of Eq. 2 accounts for the electrostatic polarization of water caused by the presence of charges of the solute. This term can be computed by solving the Poisson-Boltzmann (Honig et al., 1993; Warwicker and Watson, 1982; Sharp and Honig, 1990) or generalized-Born (Still et al., 1990) equations for a particular conformational configuration. The second term constitutes the free energy associated with the process of cavity creation for the accommodation of a hypothetical uncharged solute being transferred from vacuum to water solvent. To a good approximation, this term is linearly proportional to the total solvent-accessible surface area of the solute as follows:

$$\langle \Delta G^i_{np} \rangle = \gamma \cdot \langle SAS^i \rangle + \beta \quad (3)$$

where γ is the surface tension, β is the intercept fitting constant, and SAS is the total solvent-accessible surface area of a particular instantaneous structure of the conformer i (i.e., a snapshot saved during the MD). The values of γ and β used in this study were those previously reported for aqueous solutions 5.42 cal/mol·Å^2 and 92 cal/mol, respectively, for water (Sitkoff et al., 1994). Note that this solvation energy term ($\langle \Delta G^i_{sol} \rangle$) also accounts for the solvent entropy as averaged out in the continuum solvent.

However, the conformational energy $\langle G^i \rangle$ of a solute conformer i of Eq. 1 is given by:

$$\langle G^i \rangle = \langle E^i_{internal} \rangle - T \cdot S^i \quad (4)$$

where $\langle E^i_{internal} \rangle$ is the average internal energy in the absence of solvent and S^i is the solute conformational entropy. Generally, the internal energy term can be evaluated using either first principle quantum mechanics for small molecules (BenTal et al., 1997) or classical molecular force fields for large ones (Srinivasan et al., 1998; Jayaram et al., 1998; Lee et al., 2000, 2001), as long as the energy scale used is consistent with the method under investigation. The solute entropy term, which partly contributes to the overall system entropy along with the solvent entropy mentioned above, can be estimated using normal mode analysis (Nguyen and Case, 1985). There is currently no accurate way to compute this quantity for highly flexible molecules with a large number of degrees of freedom. However, recent works by Kollman's and Case's groups (Srinivasan et al., 1998; Lee et. al., 2000, 2001) have suggested that the contribution of this solute entropy term to the solute conformational energy difference between folded and unfolded protein states is small enough to be neglected. Nevertheless, the contribution of this entropic term to the overall solute conformational free energy difference between two states for glycoprotein will be examined in this work. Combining Eqs. 1–4, for a particular solute conformation i in aqueous solution, the absolute free energy can be approximately defined as follows:

$$G^i_{solution} \equiv \langle \Delta G^i_{elect} \rangle + (\gamma \cdot \langle SAS^i \rangle + \beta) + \langle E^i_{internal} \rangle - T \cdot S^i \quad (5)$$

Equation 5 is essentially the equation used in the MM-PBSA method proposed by Srinivasan et al. and the method used to compute the free energy difference between A-DNA and B-DBA by Jayaram et al. (1998). It has also been used to compute the binding free energy of protein-protein and protein-nucleic acid interactions, and to perform protein structural refinement (for a complete review, see Kollman et al., 2000). Kollman et al. have shown that the MM-PBSA method is a reliable method for calculating the relative free energy between two macromolecular structures in comparison with experiment.

MD simulation is normally used to generate an ensemble of conformations. Each conformation, saved as a snapshot during the course of dynamics, is then evaluated using Eq. 5 without the entropic term, and this term is later added by computing the solute entropy of a representative structure of a particular conformer. As a result, each snapshot contains a portion of the thermal energy as potential energy in the form of bond stretching, bending (vibrational modes), and van der Waals overlaps. This portion of thermal energy significantly changes the values of both internal energy and solvation free energy terms with respect to those of the corresponding minimized structures. Therefore, if α^i is the thermal energy portion, $\langle E^i_{min} \rangle$ is the average minimum internal energy of the conformation i, ω^i is the solvation free energy difference between minimized and dynamic structures, and $\langle \Delta G^i_{min} \rangle$ is the average solvation energy of the minimized structure of the conformation i, then we can obtain the following relations:

$$\langle E^i_{internal} \rangle = \langle E^i_{min} \rangle + \langle \alpha^i \rangle \quad (6)$$

$$\langle \Delta G^i_{sol} \rangle = \langle \Delta G^i_{min} \rangle + \langle \omega^i \rangle \quad (7)$$

Let

$$\Delta^i = \langle \alpha^i \rangle + \langle \omega^i \rangle \quad (8)$$

and

$$G^{i\_conf}_{solution} \equiv \langle \Delta G^i_{min} \rangle + \langle E^i_{min} \rangle - T \cdot S^i \quad (9)$$

then Eq. 5 becomes:

$$G^i_{solution} \equiv G^{i\_conf}_{solution} + \Delta^i \quad (10)$$

Equation 10 shows the relationship between the free energy of a conformer in solution taken from an MD simulation and the free energy of the minimized structure characterizing that conformer. This relationship is further explored in a later section.

SIMULATION OVERVIEW

The objective of this study is to examine the structural dynamics and to determine the lowest free energy structures of AFGP-8. The dynamics of AFGP-8 were studied by analyzing classical MD trajectories of explicitly solvated AFGP-8, and the solvent-averaged profile was constructed using absolute free energies of each AFGP-8 conformation along the trajectory path. In the MD portion of the study, the AMBER 5.0 (Case et al., 1997) program package was used with the Weiner force field (Weiner et al., 1984, 1986) for amino acid residues and water, complemented with the compatible Woods '93 (Woods et al., 1993) force field for saccharides to ensure compatible treatment of both the amino acid and sugar components of APGP-8. All MD simulations were carried out at a constant temperature (300 K) and volume with explicit inclusion of 4656 TIP3P (Jorgensen et al., 1983) water molecules and dual potential truncation applied at 15.0 Å and 20.0 Å for first and second cutoffs, respectively. Previous work (York et al., 1993) has shown that the use of dual cutoff in evaluating the electrostatic interactions gives results in good agreement with those obtained with Ewald summation techniques. Periodic boundary conditions with the minimum image convention were used for all MD simulations. A time step of 2.0 fs was used in conjunction with the SHAKE algorithm (Ryckaert et al., 1977) for restraining motion of all covalent bonds containing hydrogen atoms.

The free energy landscape for AFGP-8 was computed by calculating the absolute free energies of many MD snapshots along the trajectory path. To evaluate the absolute free energy of each snapshot, the internal energy term for AFGP-8 was computed using the same force field and parameters as used in the MD with no potential cutoff. The electrostatic term of the solvation free energy was computed by using the linear Poisson-Boltzmann equation implemented in the DELPHI-II program (Honig and Nicholls, 1995), and the nonpolar term was computed indirectly by computing total

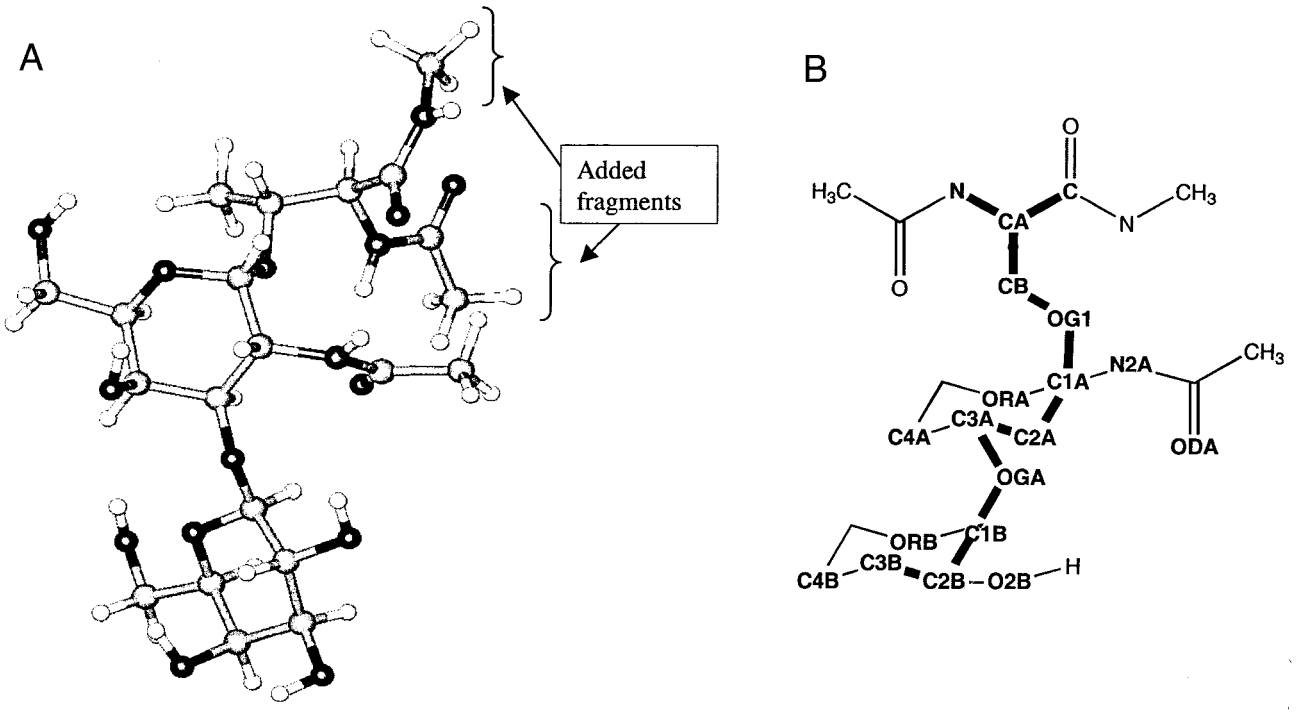

FIGURE 2 Structure of the THG residue capped by CH_3-C═O- and CH_3-NH- at the N and C termini, respectively. (*A*) The structure obtained from the quantum mechanical optimization, whereas (*B*) is a partial schematic for the capped THG with designated atom names used in the text.

solvent-accessible surface area using Sanner's algorithm implemented in the MSMS program (Sanner et al., 1996). The solute interior dielectric constant was set to 2 rather than unity, as used by Srinivasan et al. (1998). This value recognizes a modest amount of internal screening by the nonpolar groups of AFGP-8, which may be important in examining the relative energies of conformers. The solvent dielectric constant was set to 80. The dielectric boundary surface is defined by rolling a spherical probe with a radius of 1.4 Å over the AFGP-8 molecule, each atom's radius being taken from the PARSE parameter sets (Sitkoff et al., 1994).

Partial charge calculations of the THG residue

Because there are no partial charge parameters available for the THG residue, the Gaussian '98 program package (Frisch et al., 1998) was used to compute partial charges for each atom of the residue. The THG residue (Fig. 2) was built using the lowest energy conformation of each individual component of the threonine residue and two saccharides from the AMBER 5.0 library. To mimic the continuation of the protein backbone, this THG residue was capped with CH3-C(O)- and -NH-CH3 groups at N- and C-termini, respectively, as shown in Fig. 2. This capped THG residue was then optimized quantum mechanically at the restricted Hatree-Fock theory level with the 6-31G* basis set. Upon convergence, partial charges of each atom were computed using the CHELPG charge-fitting scheme (Breneman and Wiberg, 1990). These charges were used as parameters for all MD simulations and solvation free energy calculations.

The molecular dynamics

This study started with 10 NMR structures of AFGP-8 derived from the work of Lane and co-workers (1998). Each of the 10 conformers (named alphabetically from A to J) was placed in a box of 4656 water molecules with the dimensions $59.0 \times 57.0 \times 57.0$ Å^3. Minimization was carried out for 10,000 iterations using the conjugate gradient method to relax the initial simulated system. The equilibration stage was performed for 80.0 ps of MD on each simulated system under constant pressure with a harmonic restraint force constant imposed on each main chain atom (N, C_α, and C) gradually reduced from 25.0 kcal/mol·Å^2 to zero. The purpose of the restrained MD is to preserve the initial NMR structures as much as possible during this equilibration stage. During the first 50 ps, each simulated system was heated gradually from 100 to 300 K and the temperature was kept constant at 300 K thereafter. The equilibration stage was ended with 200.0 ps of MD under constant volume to bring each of the 10 simulated systems to equilibration. These 10 solvated AFGP-8 systems have the same final dimensions of $53.8 \times 51.4 \times 51.2$ Å^3.

In the production stage, each AFGP-8 system was simulated for an additional 2.0 ns of MD. During this time period, a brief thermal perturbation was introduced at the end of the first nanosecond to thermally enhance the sampling efficiency of the MD process (i.e., to sample a wider range of conformations). This thermal perturbation period lasted for 100 ps, during which the temperature was ramped from 300 to 350 K within 10 ps, sustained at this temperature for 80 ps, and ramped down to 300 K in another 10 ps. The trajectory was saved every 0.5 ps. A total of 40,000 snapshots were saved during the 20.0 ns in aggregate of MD for 10 AFGP-8 simulated systems.

Free energy

The free energy of each starting conformer of the AFGP-8 was constructed by averaging absolute free energies of each AFGP-8 conformation at every 2.0 ps along the trajectory path using Eq. 5 without the solute entropic term (the analysis of the contribution of this entropic term to the overall conformational free energy is discussed in detail below). For each snapshot, the solute was taken out of the water box and the internal energy term of Eq. 5 was computed using the same force field and parameters as used

in the MD. The electrostatic term of Eq. 5 was computed with the Delphi-II (Honig and Nicholls, 1995) program by mapping coordinates of each atom of AFGP-8 and its associated partial charge onto a cubic grid of dimension $50 \times 50 \times 50$ Å^3, whose grid resolution was 0.25 Å/grid (201^3 grid points resulted). These box dimensions were ~20–50% larger than the largest linear dimension of any AFGP-8 conformation evaluated in this work to ensure that the solute is completely enclosed inside the cubic grid. For each solvation energy calculation, 300 numerical iterations were used (as opposed to the 120 iterations typically used by the DELPHI-II program for such comparable molecule size) to ensure good convergence of the numerical results. We verified that the computed solvation energy with 300 numerical iterations is within 0.1 kcal/mol of the corresponding one with >1000 numerical iterations. Finally, the total solvent-accessible surface area term of Eq. 5 was computed by rolling a spherical probe of radius 1.4 Å over the AFGP-8 molecule using Sanner's algorithm implemented in the MSMS program (Sanner et al., 1996).

Conformation free energy

One of our two chief purposes of this study is to determine the lowest free energy structure (or set of structures) of AFGP-8 based on a set of absolute free energies of important conformations. As discussed above, because each conformation saved during the MD process intrinsically contains a portion of the thermal energy, these conformations along the trajectory path may be useful for computing the stability of a particular solute conformer (e.g., stability of A-DNA versus B-DNA over the course of dynamics) in the average sense as done by others (Srinivasan et al., 1998; Jayaram et al., 1998). However, these may not be useful for our goal in determining the lowest energy structure of the AFGP-8, because the thermal energy stored in these sampled conformations may obscure the true minimum energy surface of a relatively flexible molecule. To this end we performed another set of free energy calculations, but using minimized structures instead (computing the $G^{\text{i\_conf}}_{\text{solution}}$ term of Eq. 10). We call the energy resulting from this process the conformation free energy to distinguish it from the previous free energy, which we designate as the dynamical free energy. The procedure for computing this conformation free energy is exactly as discussed above (the dynamical free energy), but before each term of the absolute free energy (Eq. 5) is evaluated, each trajectory snapshot is subjected to minimization using conjugate gradient in an implicit solvent that provides a dielectric environment similar to the aqueous environment (a dielectric constant of $4r$, where r is the interatomic distance in Å) until the energy gradient is $<10^{-5}$ kcal/mol·Å. Because the AFGP-8 changes conformation fairly slowly, only AFGP-8 conformations at every 10.0 ps were taken for the construction of the conformation free energy profile. This approach allows us to achieve two things: 1) to select the lowest conformational free energy structure (or set of structures), and 2) to examine the relation between the $G^{\text{i\_conf}}_{\text{solution}}$ term and the $G^{\text{i}}_{\text{solution}}$ term of Eq. 10 in detail.

Estimation of solute entropy

The solute entropy term was estimated using normal mode analysis. Because this method requires minimized structures, each structure subjected to this calculation was minimized using conjugate-gradient and Newton-Raphson methods in an implicit solvent that provides a dielectric environment similar to the aqueous environment (a dielectric constant of $4r$, where r is the interatomic distance in Å) until the energy gradient is $<10^{-5}$ kcal/mol·Å. The solute entropy for each starting NMR conformer was estimated by averaging the solute entropy of four snapshots taken at equally spaced time intervals of 0.5 ns as representative solute conformations from each trajectory.

Effects of salt at the physiological pH on the free energy values of minimum structures

After the conformation free energy profile was constructed in the previous section, the 20 conformations of lowest-conformation free energy were selected and used for examining the effect of salt on the values of the absolute free energy obtained in the absence of salt. Absolute free energies of these 20 conformations were recomputed in the presence of 0.1 M 1:1 added salts (i.e., NaCl).

RESULTS AND DISCUSSION

The results of this study are presented in five parts. In the first part, we present the structural dynamics of the AFGP-8, whereas in the second part we present the dynamical and conformation free energy and lowest free-energy structures of AFGP-8 determined from the conformation free energy method. In the third part, the detailed analysis of the conformational entropy to the overall free energy is presented. Part four describes the influence of physiological salt on these low energy conformers, and part five considers intramolecular hydrogen bonds between the N-acetyl group of the carbohydrate moiety to the protein backbone to shed light on the nature of the different interpretations between the work of Dill et al. and Mimura et al. (Mimura et al., 1992; Dill et al., 1992) as discussed in the Introduction. We also present the correlation between the conformation free energy profile and the dynamical free energy discussed in part two. To help clarify the discussion that follows, the partial structure of the THG residue with our designated atom names is shown in Fig. 2 *B*.

The structural dynamics of AFGP-8

The structural dynamics of AFGP-8 can be examined by the distribution of psi-phi angles of the peptide backbone, of side chain torsional angles of each THG residue with respect to the backbone, of the torsional angles of the ether linkage of the disaccharide, and of the torsional angles characterizing carbohydrate ring conformations during the course of 20.0 ns of MD. In Fig. 3, the distribution, in the form of the contour plots, of the ψ-ϕ torsional angles of each residue of AFGP-8 from alanine 2 to proline 13 are shown. The distribution was computed from a total of 40,000 snapshots saved during the MD. It is clear from this figure that the ψ-ϕ torsional angle distributions of alanine residues 5, 8, and 11 are much broader than other residues. This implies that the AFGP-8 molecule is segmented into four major segments, pivoted at residues 5, 8, and 11 as pictorially illustrated in Fig. 4. As each individual panel of Fig. 3 shows, each residue conformation either significantly or totally adopts the conformation resembling the proline conformation ($\psi \approx 170^{\circ}$, $\phi \approx -70^{\circ}$), although residues other than proline, THG 6, 9, and 12 also adopt other conformations during the dynamics.

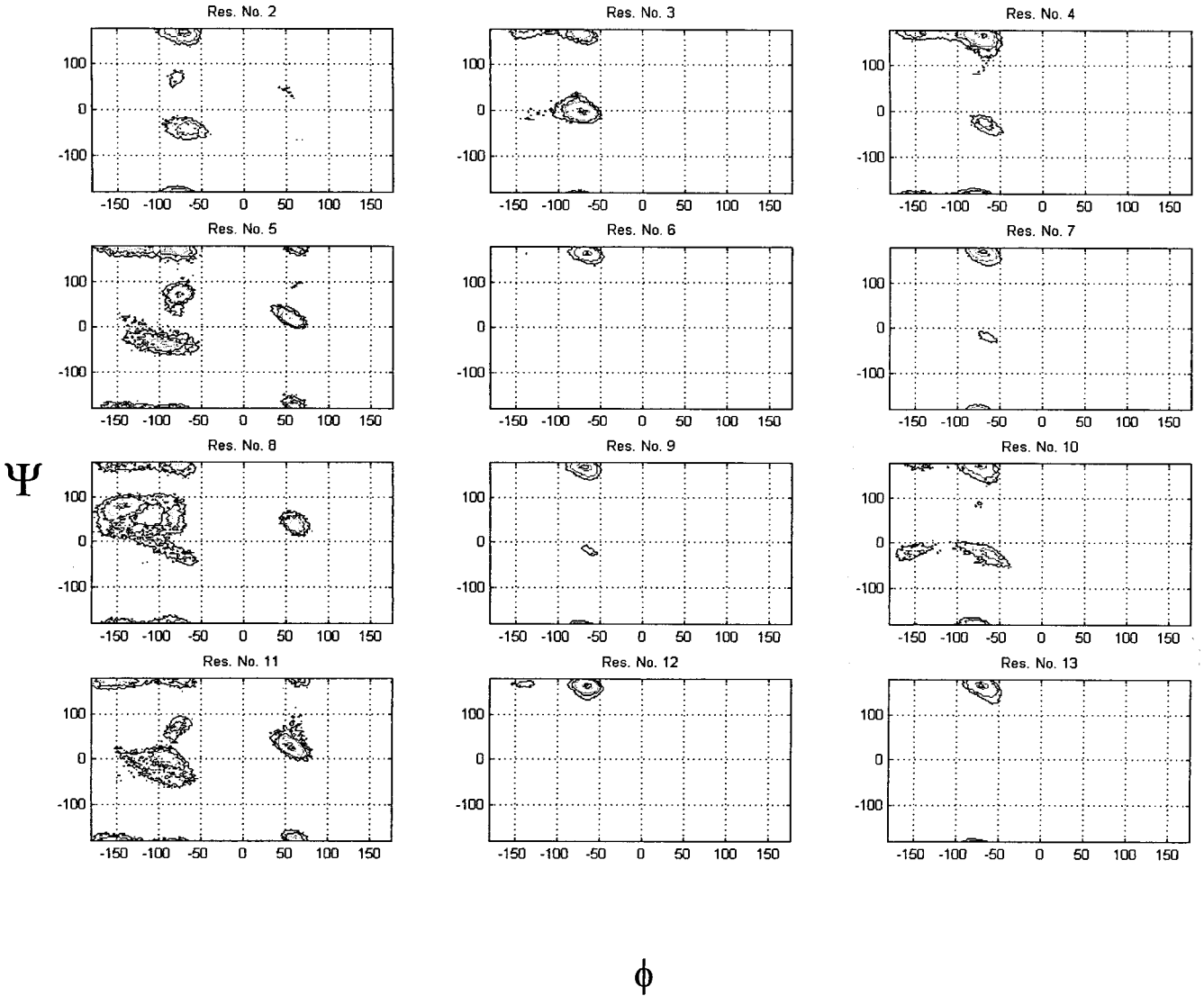

FIGURE 3 The ψ-ϕ angle distributions of each residue of AFGP-8 over the course of 20 ns of dynamics.

The torsional angles that characterize the conformation of the side chain of each THG residue with respect to the protein backbone and the carbohydrate moiety are defined as: ϕ_1(N-CA–CB-OG1), ϕ_2(CA-CB–OG1-C1A), ϕ_3(CB-OG1–C1A-C2A), ϕ_4(C4A-C3A–OGA-C1B), and ϕ_5(C3A-OGA–C1B-C2B), ϕ_6(ORA-C1A–C2A-C3A), ϕ_7(C1A-C2A–C3A-C4A), ϕ_8(ORB-C1B–C2B-C3B), and ϕ_9(C1B-C2B–C3B-C4B). These distributions are presented in Table 1. Data in this table show the preferred conformations of the disaccharide with respect to the AFGP-8 backbone during the course of the dynamics and the conformations adopted by the disaccharide ring.

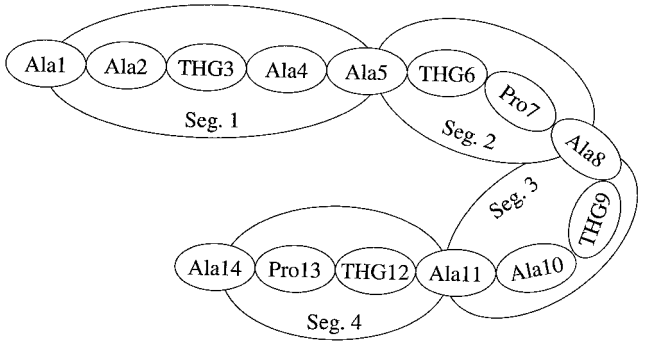

FIGURE 4 The proposed segmentation model interpreted from data in Fig. 3.

Our MD simulations of AFGP-8 suggest that the protein backbone is segmented into four segments indicated in Fig. 4, pivoted at the alanine residues 5, 8, and 11. Because other proline-containing AFGPs have proline at residues 4 and 10 instead of 7 and 13 as here, we note that segments pivoting at residues 5, 8, and 11 can equally accommodate prolines at either of position (7, 13) or (4, 10). This structural

TABLE 1 Dihedral angle distribution of the THG side chains

	THG3		THG6		THG9		THG12	
	Most Probable	Width (±)	Most Probable	Width (±)	Most Probable	Width (±)	Most Probable	Width (±)
ϕ1(N-CA–CB-OG1)	45	15	45	15	45	15	45	15
ϕ2(CA-CB–OG1-C1A)	60	25	60	25	50	25	60	15
ϕ3(CB-OG1–C1A-C2A)	170	25	170	25	170	25	170	20
ϕ4(C4A-C3A–OGA-C1B)	60	25	−60/60/160	20/25/25	60	25	60	25
ϕ5(C3A-OGA–C1B-C2B)	170	20	125/170	20/20	170	20	170	20
ϕ6(ORA-C1A–C2A-C3A)	50	15	50	10	50	10	50	10
ϕ7(C1A-C2A–C3A-C4A)	−50	7	−50	7	−50	7	−50	7
ϕ8(ORB-C1B–C2B-C3B)	55	10	55	10	55	10	55	10
ϕ9(C1B-C2B-C3B-C4B)	−55	7	−55	7	−55	7	−55	7

ϕ4 of THG6 has three most probable peaks; ϕ5 of THG6 has two most probable peaks; width is defined as the width of the distribution around the most probable peak; atom names are specified in Fig. 2 *B*.

segmentation model helps to explain experimental observations by Bush et al. and Lane et al. that the AFGP-8 molecule lacks a global, well-defined structure, because the mobility of each segment in this proposed model will effectively destroy any intramolecular long-range coupling within the AFGP-8. The segmentation model we presented here is observed for all 20 lowest free energy conformations. As shown in Fig. 9, the ranges of the ϕ angle distributions for residues 5, 8, and 11 are large compared to the remaining distributions. The most probable distributions of both ψ and ϕ from Figs. 8 and 9 generally match the corresponding distributions obtained from the MD shown in Fig. 3.

In summary, our results show that although the backbone conformation of the AFGP-8 involves flexible segments, the side chains and the disaccharides adopt well-defined conformations. This is additional evidence from simulation studies corroborating the experimental observations by Bush et al. and Lane et al. that the AFGP-8 molecule has locally defined structure, but lacks a global one.

Conformational free energy analyses

The dynamical and conformation free energies for each starting NMR conformer are shown in Fig. 5. This plot shows both the relative energy of the starting conformers and the correlation between the dynamical and conformation free energies. The energies of the starting conformers provide information on how close each NMR structure is to the lowest energy (most stable) conformation, and from this, useful structural information about them can be derived. The comparison of the dynamical and conformation energies presents the relationship between the $G^{i}_{solution}$ and $G^{i\_conf}_{solution}$ defined in Eq. 10. As shown in this figure, the conformers F and J are at the extremes of the conformational stability spectrum of the set of 10 starting conformers.

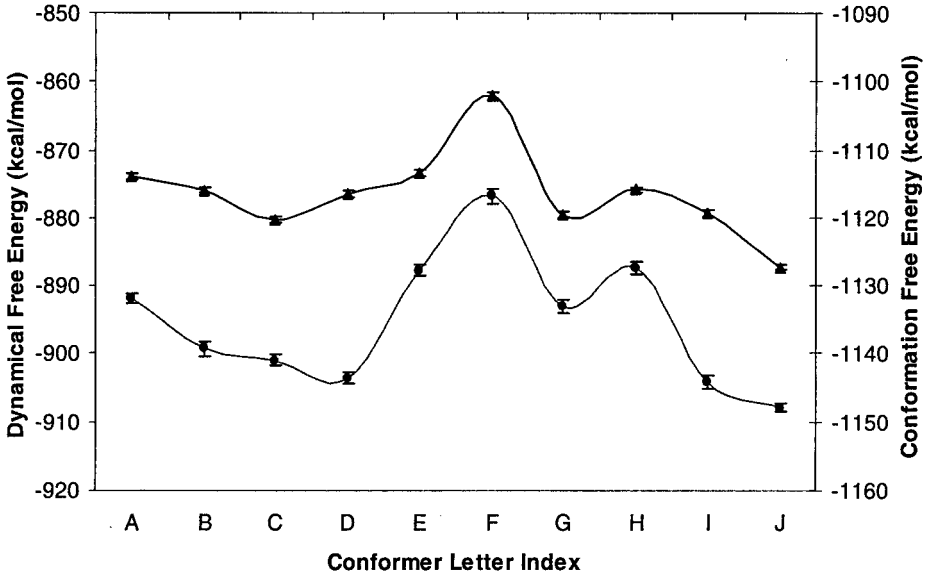

FIGURE 5 Dynamical (*triangles*) and conformation (*circles*) free energies for each starting conformer. The former are computed using Eq. 5, whereas the latter are computed using Eq. 9. These values do not contain the solute entropic term, as it is statistically the same for all 10 conformers as discussed in the text. The energy scale of the dynamical free energy is on the vertically left axis and of the conformation free energy is on the vertically right axis. Snapshots saved during the temperature ramping period are not included in computing these free energies. Each error bar represents ±one standard error.

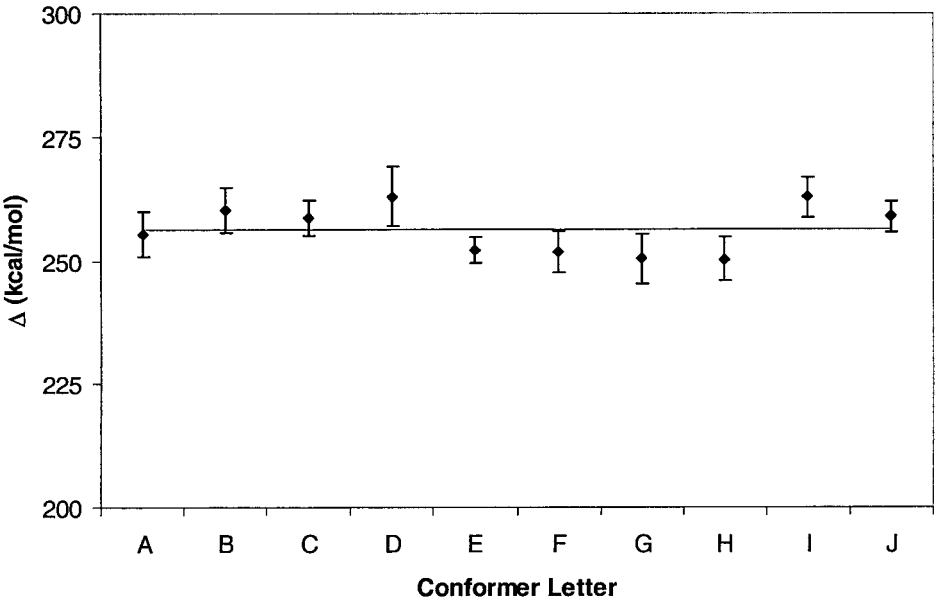

FIGURE 6 The thermal effect factor (*diamonds*) Δ as defined in Eq. 8 for each starting conformer. The error bar represents ±one standard error. The solid line is the average of these 10 Δ.

This relationship between $G^{i}_{solution}$ and $G^{i\ conf}_{solution}$, defined in Eq. 10, is analyzed further in Fig. 6, in which the values of Δ for each starting conformer are plotted. The average value of Δ is 260 kcal/mol, and its value for each starting conformer is statistically the same (within the statistical standard error, as defined to be the ratio of the standard deviation and the square root of the number of snapshots). In other words, the thermodynamics cycle presented in Fig. 1 and Eq. 1 works equivalently for both conformation free energy and dynamical free energy, because the difference of Δ^i and Δ^j, where *i* and *j* are any two starting conformers, would be statistically unimportant. This correlation establishes, quantitatively, the correlation between them as seen in Fig. 5, because the difference between them should be due to the thermal effect stored in the dynamical free energy.

Examination of both the dynamical free energies and the conformation free energies along the trajectory provide insight regarding the nature of these two different energies and the relative stabilities of the starting conformers. If the starting conformer is a stable minimum, one would expect the free energy calculations to fluctuate around a constant value as the trajectory proceeds. In contrast, the free energy should trend downward along the trajectory if the starting structure is far from a minimum (i.e., a native state of the protein) as the phase space is sampled and the channel toward the native state is traversed. We observed these two effects in the data for conformers F and J. The dynamical free energy of conformer F steadily decreased during the simulations, which is even more evident in the conformation energy plot along the trajectory. By contrast, starting conformer J exhibits nearly constant values of both dynamical free energy and conformation free energy during the simulation. This result is consistent with the fact that conformer F has highest free energy and conformer J the lowest, whether one examines either the dynamical free energy or conformation free energy.

Fig. 7 shows the histogram of the distribution of conformation free energy including data from all 10 starting NMR conformers. The bin size of the histogram is 10 kcal/mol. As shown, the lowest free energy bin contains 20 conformations. These 20 lowest free energy conformations mostly result from the starting conformers I and J, with a few from B and G. Table 2 lists these 20 conformations along with their snapshot number (at which a conformation was generated by the MD), absolute free energies sorted in increasing order, and assigned indices. We examined the standard C_α pairwise RMSD (Kabsch, 1976, 1978) of these 20 lowest free energy conformations. Of these 190 pairs, six exhibited a C_α RMSD of $<$1.5 Å, whereas the remainders were typically in the range 3.0–6.0 Å. Generally, a value $>$3.0 Å for RMSD is taken as an indicator of significant difference in conformation. Thus, the majority of these conformations are very different from each other, yet their conformation free energies are very similar. In other words, the lowest free energy conformations, while being energetically equal, are quite different in conformation. This indicates that AFGP-8 has multiple and energetically equal minima in its energy landscape. These results provide an explanation for experimental observations by Bush et al. (1981) and Lane et al. (1998) that AFGP-8 adopts many conformations in solution.

The backbone structures of the 20 lowest free energy conformations were examined. It was found that the ψ-ϕ angles tended to cluster around no more than three different values for each successive residue. The average and standard deviations for the ψ-ϕ angles in these clusters are reported in Figs. 8 and 9. The center of the symbol is placed at the average value of the angle for the cluster and the

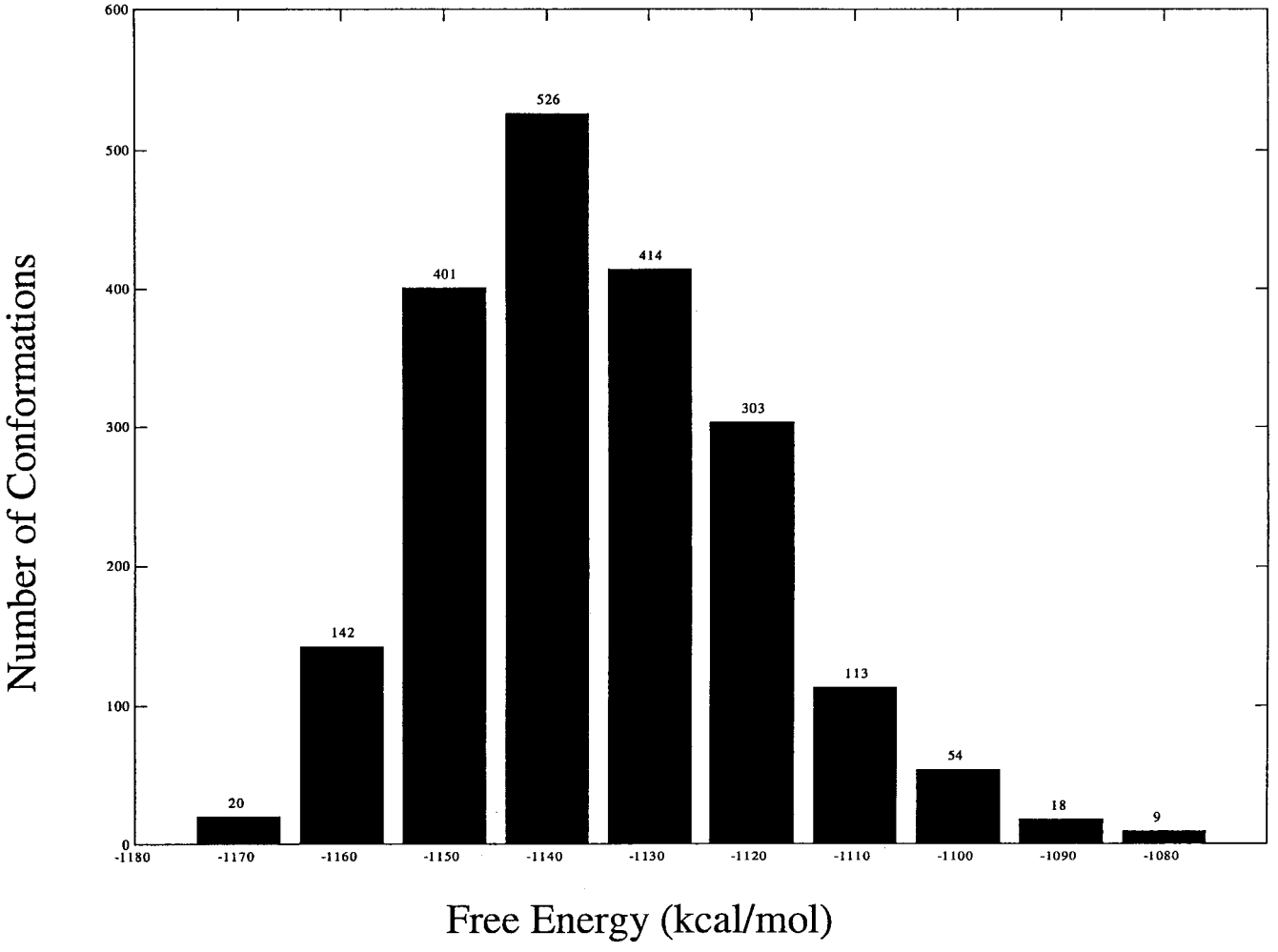

FIGURE 7 Distribution of conformation free energy.

standard deviation is plotted as an error bar. The segmentation of the backbone motion noted above for the entire set of trajectories is also found in our analysis of the 20 lowest free energy conformations. In particular, it is very clearly seen in the spread of values for the ϕ angles for residues 5, 8, and 11.

Contribution of solute entropy

Fig. 10 shows the solute entropies of each NMR starting conformer along with their standard errors. Calculation of the solute entropy contribution to the free energy (the TS^{i} term in Eqs. 5 and 9) proceeds by determining the normal modes of the energy-minimized structure and use of the resulting calculated frequencies in the statistical mechanical expression for evaluating entropy. These calculations are laborious and have not been included in the data reported in Figs. 5 and 7. The results in Fig. 10 justify this omission. Because we are interested here in the difference in conformation free energies across sampled phase space, the relative constancy of the solute entropy contribution across the variety of structures for the starting conformers indicates that the changes in this term along the conformational trajectories are small compared with the changes of the other terms in Eq. 9. This same conclusion has been reached by earlier authors (Srinivasan et al., 1998; Jayaram et al., 1998; Lee et al., 2000, 2001) using similar methods for calculation of free energy differences.

Effect of salt at the physiological pH on the conformation free energy

We explored the influence of salt in the physiologically relevant system by recalculating the absolute free energy of the 20 lowest AFGP-8 conformations in a medium of ionic strength corresponding to 0.1 M 1:1 salt. As might be expected for an uncharged molecule, the influence of salt simply shifts each free energy downward by <0.1 kcal/mol, with no effect on the relative energies of the conformers. Both the absence of differential effects of salt and the

TABLE 2 The 20 lowest free energy conformations

Indices	Conformer	Snapshot Number	Conformational Free Energy (kcal/mol)
1	I	2341	−1174.11
2	J	2741	−1173.46
3	J	3201	−1173.02
4	B	2901	−1172.25
5	J	2661	−1171.84
6	I	3981	−1168.81
7	J	2581	−1168.74
8	J	3741	−1168.46
9	J	3921	−1167.86
10	I	3961	−1167.82
11	J	0661	−1167.60
12	I	2861	−1167.14
13	I	2641	−1166.39
14	I	2021	−1165.63
15	I	1481	−1165.60
16	J	3861	−1165.48
17	I	2081	−1165.41
18	G	3501	−1165.39
19	B	1381	−1165.36
20	I	1981	−1165.28

Note that each of these conformations is given a unique index, which will be used for all later discussions and related tables.

relatively low absolute value of the shift indicate that salt has no significant role in the low-energy conformers of AFGP-8.

Role of intramolecular hydrogen bonds on stability of AFGP-8 segments

We analyzed the hydrogen-bonding patterns in the AFGP-8 configurations to shed light on the different interpretations between experimental observations by works of Mimura et al. (1992) and Dill et al. (1992) regarding the intramolecular hydrogen bond between the N-acetyl group and the protein backbone. We approached the issue by examining at which time proportions intramolecular hydrogen bonds occur during the dynamics of the most stable conformer J. In addition, because conformational stabilities of both conformers F and J are so different (~25 kcal/mol apart), we also wanted to know what role the intramolecular hydrogen bonds play in the stability of these two conformations, if there is a difference in the intramolecular hydrogen bond profiles of the two conformers. To accomplish these two goals, we tabulated the intramolecular hydrogen bond between the N-acetyl group and the protein backbone over the entire course of the dynamics for each THG residue of conformers F and J.

The carbonyl oxygen atom (ODA in Fig. 2 *b*) of the N-acetyl group is assigned to be the proton-acceptor atom, whereas the rest of the simulated system, including waters, acts as potential proton-donor atoms. A hydrogen bond is defined to exist when the distance between a proton-donor to the proton-acceptor atom is within 3.4 Å and the angle between the proton-to-donor bond and donor-to-acceptor line is within 45°. For each snapshot along the MD trajectory, the total number of hydrogen bonds between the ODA atom to those atoms that donate protons is normalized to one. Hydrogen-bonding profiles of both conformers F and J between ODA and the rest of the system for each THG residue are presented in Table 3. Summing the percentages of hydrogen bonding to the protein backbone for conformer J over each of the THR residues and dividing by 4 shows

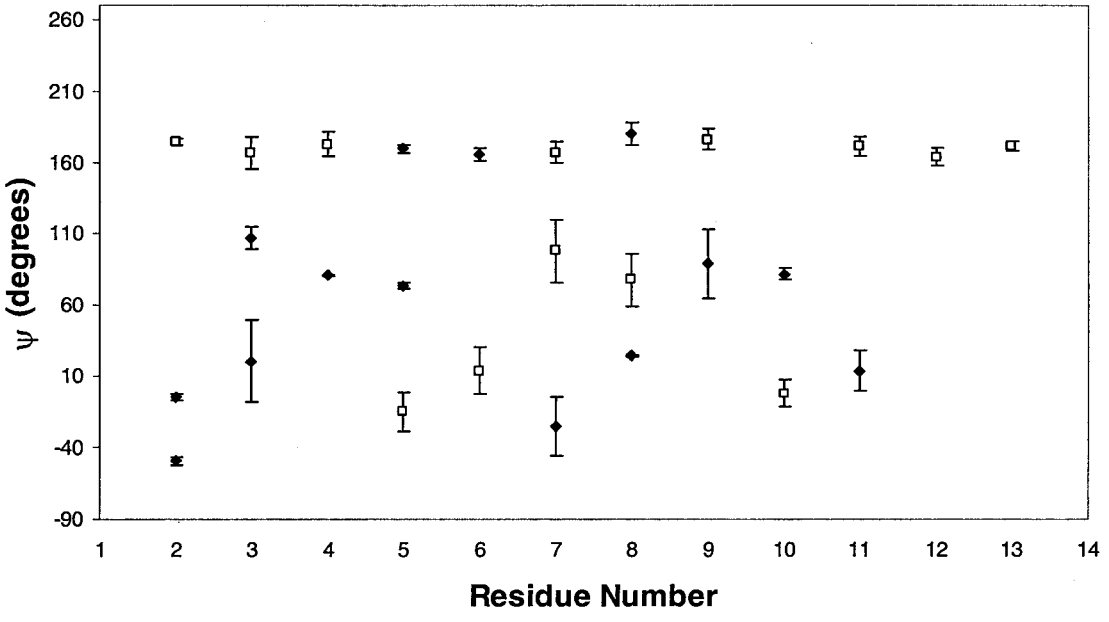

FIGURE 8 ψ-Angle distribution of the 20 lowest free energy conformations. Unfilled squares represent most probable distributions, as discussed in the text.

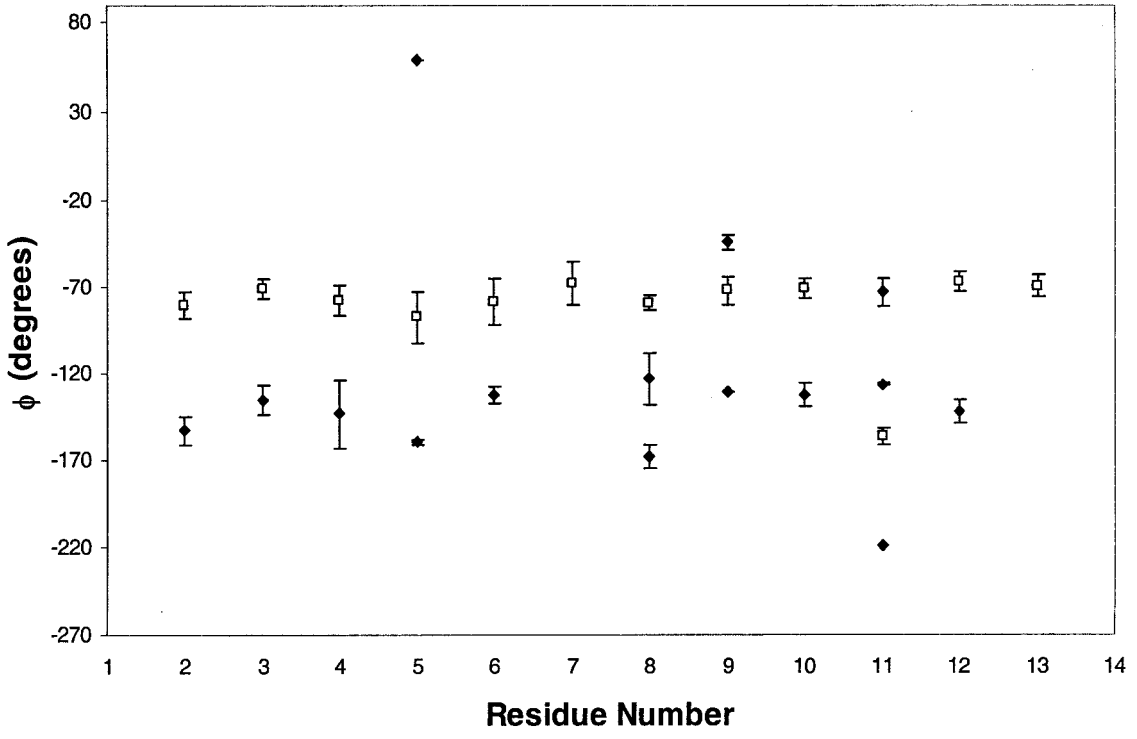

FIGURE 9 ϕ-Angle distribution of the 20 lowest free energy conformations. Unfilled squares represent most probable distributions, as discussed in the text.

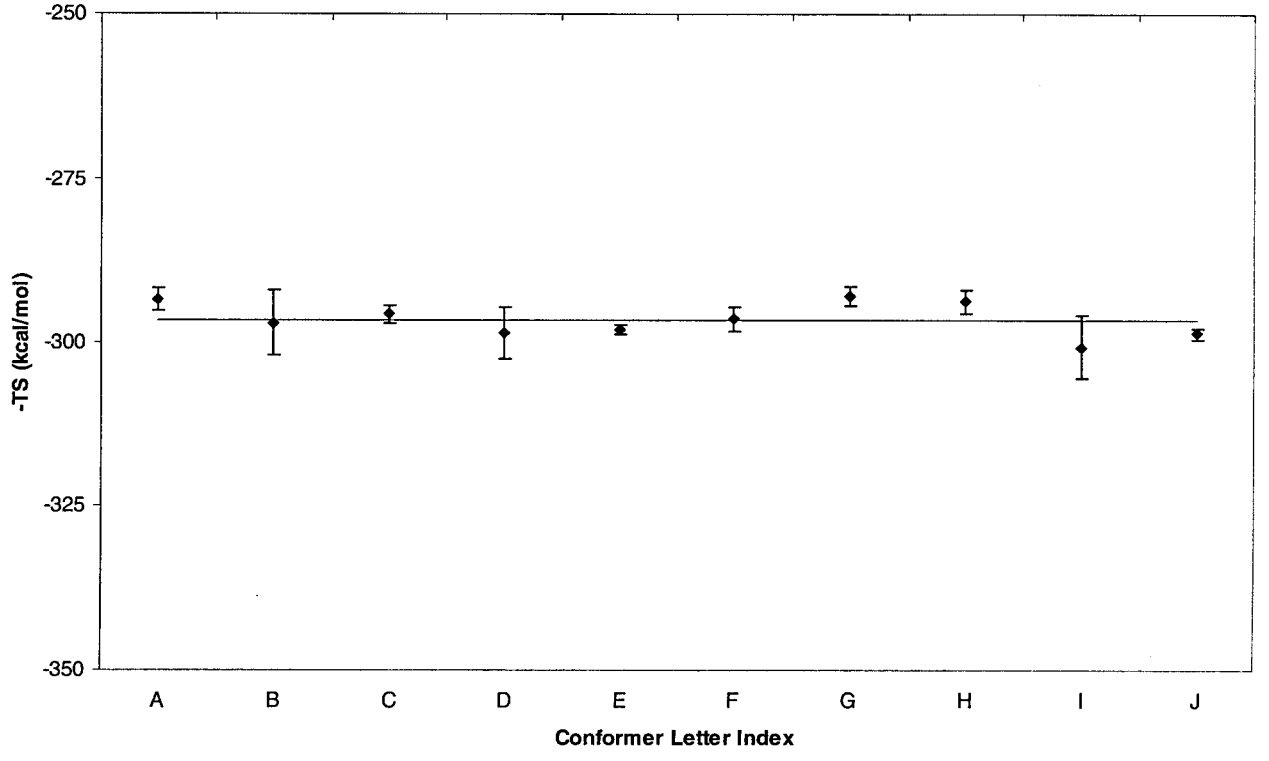

FIGURE 10 Solute entropy (*diamonds*) for each starting conformer. The solute entropy of each conformer is computed by averaging the solute entropies of four snapshots taken every 0.5 ns. The error bar represents ±one standard error from the average. The solid line is the fitting line for these solute entropies. The procedure used to compute solute entropy is discussed in the text.

TABLE 3 Intramolecular hydrogen bond profiles for both conformers F and J

	Atom Names	Conformer F	Conformer J
THG3	ALA1:N	1.2	2.8
	ALA2:N	0.0	34.6
	THG3:N	2.4	30.5
	THG3:O2B	2.2	0.0
	Water	94.1	32.2
THG6	ALA5:N	0.0	46.5
	THG6:N	1.1	49.2
	Water	98.9	4.3
THG9	THG9:N	0.0	7.5
	ALA10:N	13.3	0.0
	Water	86.7	92.5
THG12	THG12:N	9.6	0.0
	THG12:O2B	0.7	3.0
	Water	89.7	97.0

The carbonyl oxygen atom is a proton-acceptor atom. All atoms shown are proton-donor atoms. The notation shown for proton donors other than water is ⟨Residue name⟩ ⟨Residue number⟩:⟨Atom name that belongs to this particular residue⟩.

that ~43% of time there is intramolecular hydrogen bonding between the N-acetyl group and the protein backbone. Because this time proportion is <50%, it would be difficult to unambiguously identify the intramolecular hydrogen bond in the NMR spectra. Hence, it is possible that the observations by Mimura et al. and Dill et al. are reconcilable (Mimura et al., 1992; Dill et al., 1992). As for the comparison between structural stability of these two conformers, data in these two intramolecular hydrogen bond profiles suggest that, among other factors, intramolecular hydrogen bonds between the N-acetyl group and the protein backbone may contribute to the stability of AFGP-8 molecules. In addition, The intramolecular hydrogen bond may assist in the structural stability of the first segment in AFGP-8, as this segment lacks a proline residue. More generally, such intramolecular hydrogen bonds may also provide needed support for the conformational and segmental stability in other segments in other AFGP molecules.

Relevance of the simulation results to AFPG-8 function

To understand the mechanism of function of biological antifreezes or any biologically active protein, it is important to know the intrinsic structure of the protein. As such, we have studied the 20 lowest-free energy conformations resulting from the conformation free energy profile. Pairwise examination of their C_α RMSD suggests that the AFGP-8 molecule has many structurally distinct degenerate energy minima (and/or thermally accessible states) in its energy landscape. It is possible that this set of low-energy conformations plays a role in the function of AFGP-8 in retarding the growth rate of the ice crystal. While we can expect some modification of the nature of the low energy conformers when in the presence of an ice/water interface, the multiplicity of degenerate minima seems likely to persist. We suggest that the process of interconversion between the multiple minima of AFGP-8 can act as a thermal reservoir that, in effect, will retard the growth rate of the ice crystal locally where it accumulates on the ice surface. This speculation on the mechanism of action of antifreeze glycoproteins is also consistent with the observed properties of the antifreeze process, such as the hysteresis governed by the rule known as the Kelvin (or Gibbs-Thompson) relation. The main distinction between this and earlier hypotheses on the action of the antifreeze proteins is that it changes the emphasis in the mechanism from one of irreversible binding to the ice to one of local energy transfer and energy coupling. In this revised view, the ice growth rate of areas on the ice surface, where AFGP-8 accumulates, is lower than the ice growth rate of areas on the ice surface, where there is no accumulation of AFGP-8, due to the retardation effect of AFGP-8 mentioned above. This hypothesized mechanism is similar to that from our previous computational work on the winter flounder antifreeze protein (AFP) (Nguyen et al., submitted for publication). AFP has two thermally and energetically accessible states in solution: a stable dimer (the bound state) and the monomer (unbound state), and interconversion between these two states could act as a thermal reservoir.

CONCLUSIONS

We present a total of 20.0-ns dynamics of AFGP-8 and its free energy profiles, starting from NMR structures. Our simulations indicate that the AFGP-8 molecule backbone is structurally segmented into four semi-rigid segments, pivoted at alanine residues 5, 8, and 11, while its disaccharides adopt well-defined conformation with respect to the AFGP-8 backbone. This segmentation model of AFGP-8 can be generalized to the higher AFGPs in accounting for the absence of long-range order observed in NMR studies of AFGP1–5 (Lane et al., 2000).

We find that prolines have an important role in stabilizing AFGP-8 in a poly-proline II structure during a high proportion of the dynamics trajectories. This is in agreement with work of Mimura et al. (1992). We also find that intramolecular hydrogen bonds have a role in the structural stability of AFPG-8, in particular in the first segment that does not contain a proline. The intramolecular hydrogen bond profile of the carbonyl oxygen of the N-acetyl group and the protein backbone helps to reconcile different interpretations between the work of Mimura et al. and Dill et al. in the sense that the hydrogen bond would actually occur only 43% of the time. We suggest that the process of interconversion between the multiple minima of AFGP-8 can act as a thermal reservoir that, in effect, will retard the growth rate of the ice crystal locally where it accumulates on the ice surface.

In addition to the MD simulations, we computed the free energy profile of AFGP-8 by minimizing structures sampled from the MD trajectories using a continuum solvation model. These results suggest that AFGP-8 has many structurally distinct, but energetically equal, minima in its energy landscape. This important feature of AFGP-8 further supports the structural segmentation model for this small antifreeze protein. The conformation free energies calculated for different AFGP-8 conformations correlate well with the dynamical free energy methods. Therefore, the conformation free energy method can be used as an alternative method for structural refinements of a class of proteins that are highly flexible, as obtained from a variety of experimental and computational techniques such as NMR and protein structure prediction algorithms.

We thank Dr. Andrew Lane for NMR structures of AFGP-8 and Dr. Adam Zemla at Lawrence Livermore National Laboratory (LLNL) for making his LGA program, which implements the Kabsch algorithm, available to us (Zemla, A. 2000. LGA program: a method for finding 3-D similarities in protein structures. Accessed at http://predictioncenter.llnl.gov/local/lga). D.H.N. also thanks LLNL for predoctoral fellowship (SEGRF) support.

This work was carried out in part at the Lawrence Livermore National Laboratory under Contract W-7405-ENG-48 from the U.S. Department of Energy.

REFERENCES

Bashford, D., D. A. Case, C. Choi, and G. P. Gippert. 1997. Computational study of the role of solvation effects in reverse turn formation in the tetrapeptides APGD and APGN. *J. Am. Chem. Soc.* 119:4964–4971.

Ben, R. N. 2001. Antifreeze glycoproteins: preventing the growth of ice. *Chembiochem.* 2:161–166.

BenTal, N., D. Sitkoff, I. A. Topol, A. S. Yang, S. K. Burt, and B. Honig. 1997. Free energy of amide hydrogen bond formation in vacuum, in water, and in liquid alkane solution. *J. Phys. Chem. B.* 101:450–457.

Breneman, C. M., and K. B. Wiberg. 1990. Determining atom-centered monopoles from molecular electrostatic potentials. The need for high sampling density in formamide conformational analysis. *J. Comp. Chem.* 11:361–373.

Burcham, T. S., D. T. Osuga, Y. Yeh, and R. E. Feeney. 1986. A kinetic description of antifreeze glycoprotein activity. *J. Biol. Chem.* 261: 6390–6397.

Bush, C. A., R. E. Feeney, D. T. Osuga, S. Ralapati, and Y. Yeh. 1981. Antifreeze glycoprotein. Conformational model based on vacuum ultraviolet circular dichroism data. *Int. J. Peptide Protein Res.* 17:125–129.

Bush, C. A., S. Ralapati, G. M. Matson, R. B. Yamasaki, D. T. Osuga, Y. Yeh, and R. E. Feeney. 1984. Conformation of the antifreeze glycoprotein of polar fish. *Arch. Biochem. Biophys.* 232:624–631.

Case, D. A., D. A. Pearlman, J. W. Caldwell, T. E. Cheatham III, W. S. Ross, C. L. Simmerling, T. A. Darden, K. M. Merz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, D. M. Ferguson, R. J. Radmer, G. L. Seibel, U. C. Singh, P. K. Weiner, and P. A. Kollman. 1997. AMBER 5. University of California, San Francisco.

Chao, H., M. Houston, R. Hodges, C. Kay, B. Sykes, N. Loewen, P. Davies, and F. Sönnichsen. 1997. A diminished role for hydrogen bonds in antifreeze protein binding to ice. *Biochemistry.* 36:14652–14660.

DeLuca, C. I., H. Chao, F. D. Sönnichsen, B. D. Sykes, and P. L. Davies. 1996. Effect of type III antifreeze protein dilution and mutation on the growth inhibition of ice. *Biophys. J.* 71:2346–2355.

Demchuk, E., D. Bashford, G. P. Gippert, and D. A. Case. 1997. Thermodynamics of a reverse turn motif. Solvent effects and side-chain packing. *J. Mol. Biol.* 270:305–317.

Deng, G. J., D. W. Andrews, and R. A. Laursen. 1997. Amino acid sequence of a new type of antifreeze protein—from the longhorn sculpin *Myoxocephalus octodecimspinosis*. *FEBS Lett.* 402:17–20.

DeVries, A. L. 1971a. Glycoproteins as biological antifreeze agents in Antarctic fishes. *Science.* 172:1152–1155.

DeVries, A. L. 1971b. Freezing resistance in fishes. *In* Fish Physiology. W. S. Hoar and D. J. Randall, editors. Academic Press, London. 157–190.

DeVries, A. L., S. K. Komatsu, and R. E. Feeney. 1970. Chemical and physical properties of freezing point-depressing glycoproteins from Antarctic fishes. *J. Biol. Chem.* 245:2901–2908.

DeVries, A. L., J. Vandenheede, and R. E. Feeney. 1971. Primary structure of freezing point-depressing glycoproteins. *J. Biol. Chem.* 246:305–308.

Dill, K., L. H. Huang, D. W. Bearden, and R. E. Feeney. 1992. Structural studies of Antarctic fish antifreeze glycoproteins by one-dimensional and two-dimensional NMR spectroscopy. *J. Carbohydr. Chem.* 11: 499–517.

Duman, J. G., and A. L. DeVries. 1972. Freezing behavior of aqueous solutions of glycoproteins from the blood of an Antarctic fish. *Cryobiology.* 9:469–472.

Duman, J. G., and A. L. DeVries. 1974. Freezing resistance in winter flounder *Pseudopleuronectus americanus*. *Nature.* 247:237–238.

Duman, J. G., and A. L. DeVries. 1976. Isolation, characterization and physical properties of protein antifreezes from the winter flounder *Pseudopleuronectus americanus*. *Comp. Biochem. Physiol. B.* 54:375–380.

Feeney, R. E. 1974. A biological antifreeze. A glycoprotein in the blood of polar fishes lowers the freezing temperature. *American Scientist.* 62: 712–719.

Feeney, R. E., and R. Hofmann. 1973. Depression of freezing point by glycoproteins from an Antarctic fish. *Nature.* 243:357–359.

Filira, R., L. Biondi, B. Scolaro, M. T. Foffani, S. Mammi, E. Peggion, and R. Rocchi. 1990. Solid phase synthesis and conformation of sequential glycosylated polytripeptide sequences related to antifreeze glycoproteins. *Int. J. Biol. Macromol.* 12:41–49.

Fletcher, G. L., C. L. Hew, and P. L. Davies. 2001. Antifreeze proteins of teleost fishes. *Annu. Rev. Physiol.* 63:359–390.

Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, and J. A. Pople. 1998. Gaussian 98 Gaussian, Inc., Pittsburgh PA.

Harding, M., L. G. Ward, and A. D. J. Haymet. 1999. Type I "antifreeze" proteins: structure-activity studies and mechanisms of ice growth inhibition. *Eur. J. Biochem.* 264:653–665.

Haymet, A. D. J., L. G. Ward, and M. M. Harding. 1999. Winter flounder "antifreeze" proteins: synthesis and ice growth inhibition of analogues that probe the relative importance of hydrophobic and hydrogen-bonding interactions. *J. Am. Chem. Soc.* 121:941–948.

Haymet, A. D., L. G. Ward, and M. M. Harding. 2001. Hydrophobic analogues of the winter flounder "antifreeze" protein. *FEBS Lett.* 491: 285–288.

Haymet, A. D. J., L. G. Ward, M. M. Harding, and C. A. Knight. 1998. Valine substituted winter flounder "antifreeze": preservation of ice growth hysteresis. *FEBS Lett.* 430:301–306.

Hew, C. L., S. Joshi, N.-C. Wang, M.-H. Kao, and V. S. Ananthanarayanan. 1985. Structures of shorthorn sculpin antifreeze polypeptides. *Eur. J. Biochem.* 151:167–172.

Honig, B., and A. Nicholls. 1995. Classical electrostatics in biology and chemistry. *Science.* 268:1144–1149.

Honig, B., K. Sharp, and A. S. Yang. 1993. Macroscopic models of aqueous solutions: biological and chemical applications. *J. Phys. Chem.* 97:1101–1109.

Jayaram, B., D. Sprous, M. A. Young, and D. L. Beveridge. 1998. Free energy analysis of the conformational preferences of A and B forms of DNA in solution. *J. Am. Chem. Soc.* 120:10629–10633.

Jia, Z. C., C. I. DeLuca, and P. L. Davies. 1995. Crystallization and preliminary x-ray crystallographic studies on type III antifreeze protein. *Protein Sci.* 4:1236–1238.

Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. 1983. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79:926–935.

Kabsch, W. 1976. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A.* 32:922–923.

Kabsch, W. 1978. A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. A.* 34:827–828.

Kollman, P. A., I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham. 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33:889–897.

Komatsu, S. K., A. L. DeVries, and R. E. Feeney. 1970. Studies of the structure of freezing point-depressing glycoproteins from an Antarctic fish. *J. Biol. Chem.* 245:2909–2913.

Kuroda, T. 1991. Proc. 4th Topical Conference on Crystal Growth Mechanisms. Hokkaido Press, Japan. p. 157.

Lane, A. N., L. M. Hays, R. E. Feeney, L. M. Crowe, and J. H. Crowe. 1998. Conformational and dynamic properties of a 14 residue antifreeze glycopeptide from Antarctic cod. *Protein Sci.* 7:1555–1563.

Lane, A. N., L. M. Hays, N. Tsvetkova, R. E. Feeney, L. M. Crowe, and J. H. Crowe. 2000. Comparison of the solution conformation and dynamics of antifreeze glycoproteins from Antarctic fish. *Biophys. J.* 78:3195–3207.

Lee, M. R., D. Baker, and P. A. Kollman. 2001. 2.1 and 1.8 angstrom average C-alpha RMSD structure predictions on two small proteins, HP-36 and S15. *J. Am. Chem. Soc.* 123:1040–1046.

Lee, M. R., Y. Duan, and P. A. Kollman. 2000. Use of MM-PB/SA in estimating the free energies of proteins: application to native, intermediates, and unfolded villin headpiece. *Proteins.* 39:309–316.

Liou, Y. C., A. Tocilj, P. L. Davies, and Z. C. Jia. 2000. Mimicry of ice structure by surface hydroxyls and water of a beta-helix antifreeze protein. *Nature.* 406:322–324.

Mimura, Y., Y. Yamamoto, Y. Inoue, and R. Chujo. 1992. NMR study of interaction between sugar and peptide moieties in mucin-type model glycopeptides. *Int. J. Biol. Macromol.* 14:242–248.

Mulvihill, D. M., K. F. Geoghegan, Y. Yeh, K. DeRemer, D. T. Osuga, F. C. Ward, and R. E. Feeney. 1980. Antifreeze glycoproteins from Polar fish. Effects of freezing conditions on cooperative function. *J. Biol. Chem.* 255:659–662.

Ng, N. F. L., and C. L. Hew. 1992. Structure of an antifreeze polypeptide from the sea raven: disulfide bonds and similarity to lectin-binding proteins. *J. Biol. Chem.* 267:16069–16075.

Ng, N. F., K.-Y. Trinh, and C. L. Hew. 1986. Structure of an antifreeze polypeptide precursor from the sea raven, *Hemitripterus americanus*. *J. Biol. Chem.* 261:15690–15695.

Nguyen, D. T., and D. A. Case. 1985. On finding stationary states on large-molecule potential energy surfaces. *J. Phys. Chem.* 89: 4020–4026.

Rao, B. N. N., and C. A. Bush. 1987. Comparison by proton NMR spectroscopy of the conformation of the 2600 dalton antifreeze glycopeptide of polar cod with that of the high molecular weight antifreeze glycoproteins. *Biopolymers.* 26:1227–1244.

Ryckaert, J. P., G. Ciccotti, and H. J. C. Berendsen. 1977. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comp. Phys.* 23: 327–341.

Sanner, M. F., J. C. Spehner, and A. J. Olson. 1996. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers.* 38:305–320.

Sharp, K. A., and B. Honig. 1990. Electrostatic interactions in macromolecules—theory and applications. *Annu. Rev. Biophys. Biophys. Chem.* 19:301–332.

Sitkoff, D., D. J. Lockhart, K. A. Sharp, and B. Honig. 1994. Calculation of electrostatic effects at the amino terminus of an alpha helix. *Biophys. J.* 67:2251–2260.

Slaughter, D., G. L. Fletcher, V. S. Ananthanarayanan, and C. L. Hew. 1981. Antifreeze proteins from the sea raven, *Hemitripterus americanus*: further evidence for diversity among fish polypeptide antifreezes. *J. Biol. Chem.* 256:2022–2026.

Smith, K. C., and B. Honig. 1994. Evaluation of the conformational free energies of loops in proteins. *Proteins: Struct., Funct., Genet.* 18: 119–132.

Sönnichsen, F. D., C. I. DeLuca, P. L. Davies, and B. D. Sykes. 1996. Refined solution structure of type III antifreeze protein-hydrophobic groups may be involved in the energetics of the protein-ice interaction. *Structure.* 4:1325–1337.

Srinivasan, J., T. E. Cheatham, P. Cieplak, P. A. Kollman, and D. A. Case. 1998. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices. *J. Am. Chem. Soc.* 120:9401–9409.

Still, W. C., A. Tempczyk, R. C. Hawley, and T. Hendrickson. 1990. Semianalytical treatment of solvation for molecular mechanics and dynamics. *J. Am. Chem. Soc.* 112:6127–6129.

Warwicker, J., and H. C. Watson. 1982. Calculation of the electric potential in the active site cleft due to alpha-helix dipoles. *J. Mol. Biol.* 157: 671–679.

Weiner, S. J., P. A. Kollman, D. A. Case, U. Chandra Singh, C. Ghio, G. Alagona, S. Profeta, Jr., and P. Weiner. 1984. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.

Weiner, S. J., P. A. Kollman, D. T. Nguyen, and D. A. Case. 1986. An all atom force field for simulations of proteins and nucleic acids. *J. Comp. Chem.* 7:230–252.

Woods, R. J., R. A. Dwek, C. J. Edge, and B. Fraserreid. 1993. Molecular mechanical and molecular dynamical simulations of glycoproteins and oligosaccharides. 1. Glycam-93 parameter development. *J. Phys. Chem.* 99:3832–3846.

Yang, A. S., B. Hitz, and B. Honig. 1996. Free energy determinants of secondary structure formation. 3. Beta-turns and their role in protein folding. *J. Mol. Biol.* 259:873–882.

Yeh, Y., and R. E. Feeney. 1996. Antifreeze proteins: structures and mechanism of functions. *Chem. Rev.* 96:601–617.

York, D. M., T. A. Darden, and L. G. Pedersen. 1993. The effect of long-range electrostatic interactions in simulations of macromolecular crystals—a comparison of the Ewald and truncated list methods. *J. Chem. Phys.* 99:8345–8348.